



# Measurement of the Top Quark Mass using a Matrix Element Method in a $b$ -Tagged Dilepton Sample

The CDF Collaboration  
URL <http://www-cdf.fnal.gov>  
(Dated: July 24, 2006)

We report a measurement of the top quark mass using events collected by the CDF II Detector from  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV at the Fermilab Tevatron. We calculate a likelihood function for the top mass in events that are consistent with  $t\bar{t} \rightarrow \bar{b}\ell^-\bar{\nu}_\ell b\ell'^+\nu'_\ell$  decays. We use events in which at least jet is identified as coming from a  $b$ -quark. The likelihood is formed as the convolution of the leading-order matrix element and detector resolution functions. The joint likelihood is the product of likelihoods for each of 30 events collected in  $955 \text{ pb}^{-1}$  of integrated luminosity, yielding a top quark mass,  $M_t = 167.3 \pm 4.6(\text{stat.}) \pm 3.8(\text{syst.}) \text{ GeV}/c^2$ .

## I. INTRODUCTION

Precision measurements of the top quark mass,  $M_t$ , place constraints on the masses of particles to which the top quark contributes radiative corrections, including the unobserved Higgs boson and particles in extensions to the standard model. At the Tevatron, top quarks are primarily produced in pairs. The dilepton channel, consisting of the decays  $t\bar{t} \rightarrow \bar{b}\ell^-\bar{\nu}_\ell b\ell'^+\nu'_\ell$ , has a small branching fraction but allows measurements which are less reliant on the calibration of the jet energy scale, the dominant systematic uncertainty, than measurements in channels with hadronic  $W$  decays. A discrepancy from other channels could indicate contributions from new processes.

The reconstruction of the top mass from dilepton events poses a particular challenge as the two neutrinos from  $W$  decays are undetected. Previous measurements in this channel [1, 2] using Run I data have calculated the mass by making several kinematic assumptions and integrating over the remaining unconstrained quantity. To extract maximum information from the small sample of dilepton events, we adapt a technique pioneered for the analysis of  $t\bar{t} \rightarrow b\ell\nu_\ell\bar{b}q\bar{q}'$  decays [3–5]. This technique uses the leading-order production cross-section and a parameterized description of the jet energy resolution. Making minimal kinematic assumptions and integrating over six unconstrained quantities, we obtain per-event likelihoods in top mass which can be directly multiplied to obtain the joint likelihood from which  $M_t$  is determined. This method was first applied to dilepton events at CDF using  $340 \text{ pb}^{-1}$  of data [7, 8].

We have applied this method to dilepton events at CDF [6] using  $1 \text{ fb}^{-1}$  of data yielding the single most precise measurement of the top quark mass in this channel. In this note we report a measurement using the same method

in a sample whose purity has been increased by requiring that at least one jet in each event be identified as resulting from a  $b$ -quark.

## II. EVENT SELECTION

This analysis is based on an integrated luminosity of  $955 \text{ pb}^{-1}$  collected with the CDF II detector between March 2002 and March 2006. The CDF II detector is a general purpose detector described elsewhere [10]. For this analysis, we select events with two high- $p_T$  leptons, missing transverse energy ( $\cancel{E}_T$ ) and two energetic jets coming from the hadronization of the  $b$ -quarks. We use the selection described as “DIL” in [12] to measure the cross-section in the dilepton channel.

The data are collected with an inclusive lepton trigger that requires events to have a lepton with  $E_T > 18 \text{ GeV}$  (for an electron) or  $p_T > 18 \text{ GeV}$  (for a muon). After full event reconstruction we require events with two leptons, both with  $E_T > 20 \text{ GeV}$  ( $p_T > 20 \text{ GeV}$  for muons) and at least one of which is isolated [11]. Candidate events must have at least two jets with  $E_T > 15 \text{ GeV}$  and be measured within  $|\eta| < 2.5$ . We also require candidate events to have  $\cancel{E}_T > 25 \text{ GeV}$  and in events with  $\cancel{E}_T < 50 \text{ GeV}$  that the  $\cancel{E}_T$  vector is at least  $20^\circ$  from the closest lepton or jet. Additionally, we require at least one jet in each event be identified as resulting from a  $b$ -quark via secondary vertex tagging. Secondary vertex tagging identifies tracks associated with a displaced secondary vertex.

## III. METHOD

The information contained in an event regarding the top mass can be expressed as the conditional probability  $P(\mathbf{x}|M_t)$ , where  $M_t$  is the top pole mass and  $\mathbf{x}$  is a vector of measured event quantities. We calculate the posterior probability using the theoretical description of the  $t\bar{t}$  production process expressed with respect to the measured event quantities:

$$P(\mathbf{x}|M_t) = \frac{1}{\sigma(M_t)} \frac{d\sigma(M_t)}{d\mathbf{x}}$$

where  $\frac{d\sigma}{d\mathbf{x}}$  is the per-event differential cross-section.

To evaluate the probability, we integrate over quantities which are unknown because they are unmeasured by the detector, such as neutrino energies. Quark energies are not directly measured, but are estimated from the observed energies of the corresponding jets. We parametrize this uncertainty using a transfer function between quark and jet energies,  $W(p, j)$ , giving us the probability of measuring jet energy  $j$  given parton energy  $p$ . We form the transfer function by fitting a double Gaussian to a predicted distribution of parton-jet energy difference from simulated events. The transverse momentum of the  $t\bar{t}$  system can also not be directly measured, but can be inferred from the unclustered energy in the event and from jets resulting from radiation. We parameterize the relation between these measured quantities and the  $p_T$  of the  $t\bar{t}$  system as a transfer function,  $W(p_T, U)$ . The total expression for the probability of a given pole mass for a specific event can be written as

$$P(\mathbf{x}|M_t) = \frac{1}{N} \int d\Phi_8 |\mathcal{M}_{t\bar{t}}(p; M_t)|^2 \prod_{jets} W(p, j) W(p_T, U) f_{PDF}(q_1) f_{PDF}(q_2) \quad (1)$$

where the integral is over the entire six-particle phase space,  $q$  is the vector of incoming parton-level quantities,  $p$  is the vector of resulting parton-level quantities: lepton and quark momenta, and  $|\mathcal{M}_{t\bar{t}}(p; M_t)|$  is the  $t\bar{t}$  production matrix element as defined in [14, 15]. The constant term in front of the integral ensures that the normalization condition for the probability:

$$\int d\mathbf{x} P(\mathbf{x}|M_t) = 1$$

is satisfied.

### A. Background

The probability  $P(\mathbf{x}|M_t)$  is sufficient to extract the top quark mass in an unpolluted sample. However, the top quark candidate events collected by CDF have a small fraction of background events which mimic the top quark

signature. To reduce the effect of these events on the measurement, we calculate the probabilities,  $P_{bg_i}(\mathbf{x})$  that they were produced by a given background process; we form the generalized per-event probability as

$$P(\mathbf{x}|M_t) = P_s(\mathbf{x}|M_t)p_s + P_{bg_1}(\mathbf{x})p_{bg_1} + P_{bg_2}(\mathbf{x})p_{bg_2} \dots \quad (2)$$

simply a sum of the probabilities for each process, weighted by their respective priors. Here,  $P_s(\mathbf{x}|M_t)$  is as described in equation 1 and the  $P_{bg_i}(\mathbf{x})$  are formed by calculating a differential cross-section for each event in a manner similar to  $t\bar{t}$ . The background processes for which we evaluate probabilities for in this manner are: Drell-Yan with associated jets and  $W+3$  jets production where one jet is incorrectly identified as a lepton.

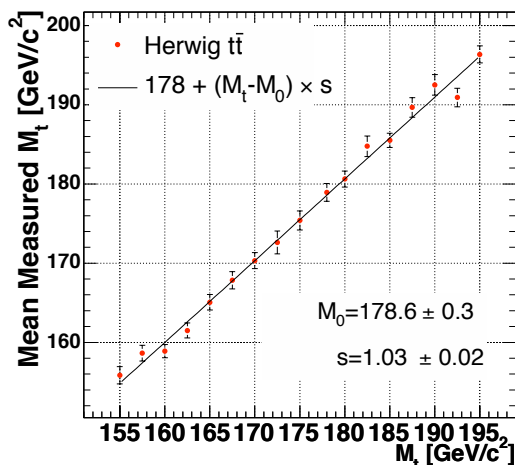
The weights for each term in Equation 2 are determined in part from the number of expected background events in each category [13]. These numbers are listed in Table I.

Source	$N_{ev}$
$WW$	$0.0 \pm 0.1$
$WZ$	$0.0 \pm 0.1$
$Z/\gamma^* \rightarrow ee, \mu\mu$	$0.38^{+0.76}_{-0.38}$
$Z/\gamma^* \rightarrow \tau\tau$	$0.0 \pm 0.1$
Fakes	$0.81^{+1.62}_{-0.81}$
$t\bar{t}$ ( $M_t = 175 \text{ GeV}/c^2$ )	$32.4 \pm 1.3$
Total SM expectation	$33.6 \pm 3.2$
Data ( $\int \mathcal{L} dt = 955 \text{ pb}^{-1}$ )	30

**TABLE I:** Expected signal and background events and their sources for a data sample of  $\int \mathcal{L} dt = 955 \text{ pb}^{-1}$ .

#### IV. CALIBRATION

To test the performance of the method, we construct Monte Carlo experiments using Monte Carlo for generated top masses from  $155 \text{ GeV}/c^2$  to  $195 \text{ GeV}/c^2$ . The number of signal and background events in each Monte Carlo experiment are Poisson fluctuated values around the *a priori* estimates given in Table I; the estimate for  $t\bar{t}$  at varying masses is evolved to account for the variation of cross-section and acceptance. The response of the method for Monte Carlo experiments with both signal and background is shown in Figure 1. A correction, as derived from this response, is applied to the measured value in data.



**FIG. 1:** Response for Monte Carlo experiments of signal and background events.

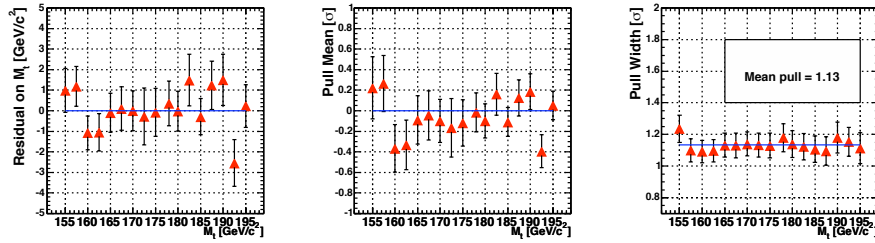


FIG. 2: Residual, pull mean, and pull width for varying top mass Monte Carlo samples.

In the interest of computational tractability, several assumptions are made in the evaluation of the integrals in Equation 1, such as the leading two jets in the event coming from  $b$ -quarks from top decay, and lepton momenta and jet angles being measured perfectly. These assumptions are violated in small and understood ways in realistic events. Due to these effects, the method underestimates the statistical error. The pull distribution in Monte Carlo experiments is shown in 2. To account for this underestimation, we scale the statistical error by a factor,  $S = 1.13$ , derived from the results of our Monte Carlo experiments.

## V. SYSTEMATIC UNCERTAINTIES

There are several sources of systematic error in our measurement which are summarized in Table II.

Source	Size ( $\text{GeV}/c^2$ )
Jet Energy Scale	3.3
Multiple Interactions	0.2
Lepton Energy Scale	0.1
Generator	0.7
Method	0.3
Sample composition uncertainty	1.0
Background MC	0.8
Background modeling	0.4
B-jet tagging	0.3
FSR modeling	0.4
ISR modeling	0.4
PDFs	0.8
Total	3.8

TABLE II: Summary of systematic errors.

The single largest source of systematic error comes from the uncertainty in the jet energy scale, which we estimate by varying the jet energy corrections by  $\pm 1\sigma$  and is  $3.3 \text{ GeV}/c^2$ . We assign an uncertainty to the presence of multiple interactions per beam crossing during high-luminosity; this uncertainty is measured to be  $0.2 \text{ GeV}/c^2$ . A small uncertainty is assigned to possible imperfections in the modeling of lepton energy in Monte Carlo events; this uncertainty is measured to be  $0.1 \text{ GeV}/c^2$ . The uncertainty in the Monte Carlo generator used to perform Monte Carlo experiments, estimated by measuring the difference in extracted the top mass from PYTHIA events and HERWIG events, amounts to  $0.7 \text{ GeV}/c^2$ . The uncertainty in the response correction shown in Figure 1 is estimated by varying that response by  $\pm 1\sigma$  and is  $0.3 \text{ GeV}/c^2$ . The uncertainty due to initial-state (ISR) and final-state (FSR) radiation is estimated by varying the amount of ISR and FSR in simulated events and is measured to be  $0.4 \text{ GeV}/c^2$  for both cases.

The uncertainty in background composition is estimated by varying the background estimates from Table I within their errors and amounts to  $1.0 \text{ GeV}/c^2$ . In addition, a large uncertainty comes from the limited number of Monte Carlo background events available for Monte Carlo experiments. To measure this uncertainty, we split each background sample into twenty pairs of disjoint sets. We measure the mass for each of the disjoint sets and take the RMS of the difference between them as an estimate of the error. Summing these, we get  $0.8 \text{ GeV}/c^2$ . We estimate an

uncertainty coming from possible imperfections in modeling the two largest sources of background: Drell-Yan and events with a “fake” lepton. This uncertainty is estimated to be  $0.4 \text{ GeV}/c^2$ . In our pseudo-experiments, we utilize data events for the fake lepton background. Due to limited statistics, we do not require  $b$ -jet identification on these events. To estimate the uncertainty due to these events, we alternatively use Monte Carlo events in which we require  $b$ -jet identification. We take the difference between the measured mass using these two pools of fake events to be a systematic uncertainty. We measure this uncertainty to be  $0.4 \text{ GeV}/c^2$ .

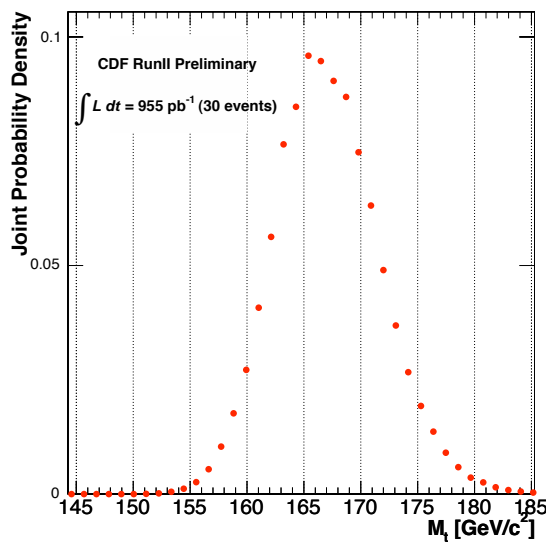
Finally, the uncertainties in the parton distribution functions (PDFs) are estimated by using different PDF sets (CTEQ5L vs. MRST72), different values of  $\Lambda_{QCD}$ , and varying the eigenvectors of the CTEQ6M set, yielding a total uncertainty of  $0.8 \text{ GeV}/c^2$ .

## VI. RESULT IN DATA

We apply the procedure described in Section III to the 30 candidate events observed in the data. After applying the corrections described in Section IV, we measure a top quark mass of

$$M_{top} = 167.3 \pm 4.6(\text{stat.}) \text{ GeV}/c^2$$

The final posterior probability density for the events in data can be seen in Figure 3.



**FIG. 3:** Final posterior probability density as a function of the top pole mass for the 30 candidate events in data.

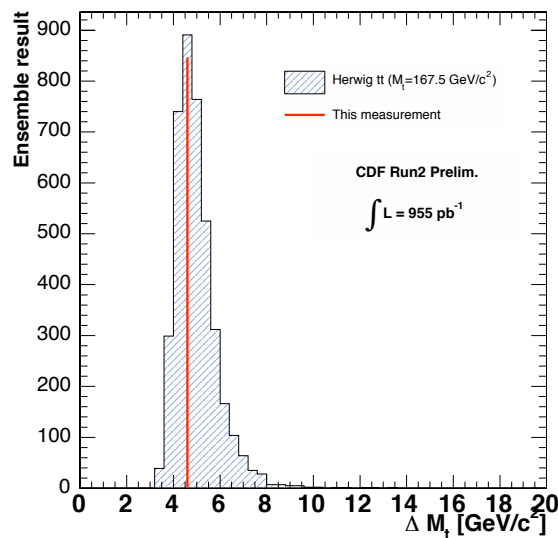
The measured statistical uncertainty is consistent with that measured for Monte Carlo experiments using  $M_{top} = 165 \text{ GeV}/c^2$  signal events (which had a mean *a priori* statistical uncertainty of  $5.0 \text{ GeV}/c^2$ ) as shown in Figure 4.

## VII. CONCLUSION

We measure the top quark mass to be

$$M_{top} = 167.3 \pm 4.6(\text{stat.}) \pm 3.8(\text{syst.}) \text{ GeV}/c^2$$

in dilepton events requiring at least one identified  $b$ -jet in  $955 \text{ pb}^{-1}$  of CDF II data. We have used a normalized per-event differential cross-section for leading order top quark pair production and background to form a posterior probability. The statistical power of this method allows this measurement to have comparable uncertainty to the measurement made in a sample not requiring  $b$ -jet identification, despite having nearly twice the available statistics.



**FIG. 4:** Distribution of expected errors for  $M_t = 167.5 \text{ GeV}/c^2$ . The measured error is shown as the line; 39% of Monte Carlo experiments yielded a smaller error.

### Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comisión Interministerial de Ciencia y Tecnología, Spain; in part by the European Community's Human Potential Programme under contract HPRN-CT-2002-00292; and the Academy of Finland.

- 
- [1] B. Abbot *et al.*, (DØ Collaboration), Phys. Rev. Lett. **80**, 2063 (1998).
  - [2] F. Abe *et al.*, (CDF Collaboration), Phys. Rev. Lett. **82**, 271 (1999).
  - [3] K. Kondo, J. Phys. Soc. Jpn. **57**, 4126 (1988).
  - [4] J. Estrada, FERMILAB-THESIS-2001-07 (2001).
  - [5] F. Canelli, FERMILAB-THESIS-2003-22 (2003).
  - [6] The CDF Collaboration, CDF/TOP/PUB/8369.
  - [7] A. Abulencia *et al.*, The CDF Collaboration, Phys. Rev. Lett. **96**, 152002 (2006).
  - [8] A. Abulencia *et al.*, The CDF Collaboration, To be published in Phys. Rev. D., hep-ex/0605118.
  - [9] The CDF Collaboration, CDF/TOP/PUB/8090.
  - [10] D. Acosta, et al., The CDF Collaboration, Phys. Rev. D **71**, 032001 (2005).
  - [11] An isolated lepton is one for which no more than 10% extra energy is measured in a cone of  $\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \leq 0.4$  around the lepton.
  - [12] D. Acosta et al., The CDF Collaboration, Phys. Rev. Lett. **93**, 142001 (2004).
  - [13] The CDF Collaboration, CDF/TOP/PUB/8103.
  - [14] G. Mahlon, S. Parke, Phys. Lett. B **411**, 173 (1997).
  - [15] G. Mahlone, S. Parke, Phys. Rev. D **55**, 7249 (1996).